

Performance Analysis of Stereo Matching Using Segmentation Based Disparity Map

Arti Khaparde, Apurva Naik, Manini Deshpande, Sakshi Khar, Kshitija Pandhari, Mayura Shewale

Department of Electronics and Telecommunication

Maharashtra Institute of Technology

Pune, Maharashtra, India

artikhaparde@gmail.com, apurva.naik@gmail.com, deshpane.manini@gmail.com

Abstract— Stereo vision has been studied extensively due to its usefulness in many applications like 3D scene reconstruction, robot navigation, etc. Rather than finding out the disparity between two original stereo images, various segmentation techniques are used to segment the images and the disparity between the resulting segmented images is calculated. The comparison between the disparity of the original stereo image pair and that of the segmented image pair is done on the basis of compression ratio and Peak-Signal to Noise Ratio (PSNR), which is calculated for image quality measurement. Segmentation techniques like Mean Shift Algorithm, K-means Algorithm and Particle Swarm Optimization (PSO) are used and their results are compared on the basis of subjective and objective parameters. The experimental results show that PSO based 3D image reconstruction gives a good compromise between subjective quality and compression ratio.

Keywords-Disparity Map; PSNR; Mean Square Error (MSE); Compression ratio; Particle Swarm Optimization.

I. INTRODUCTION

Stereo Image matching is one of the core research areas in Computer Vision and Digital Photogrammetry. Technological developments in stereo image matching have advanced from the primitive area based cross-correlation technique to more and more precise feature-based and area-based matching [9]. Stereo allows us to recover information from the given two images about a three dimensional location of objects, which does not exist in any single image. The main goal of stereo image matching is to recover depth information from the given two or multiple images. In order to recover depth information the stereo images should be brought into point-point correspondence. Correspondence points are the projections of a single point into the three dimensional scene. The difference between the locations of these two correspondence points is known as parallax or disparity, which is a function of position of the point in the scene, orientation and physical characteristics of the camera. So, disparity can be used as constraint for matching. Although, feature-based techniques are more accurate, but they generate sparse disparity maps [1, 2, 4, 8]. Hence, in this paper, we are proposing an area-based segmentation method, as it generates a dense disparity map for 3D reconstruction.

The paper is organized as follows: Section 1 gives brief introduction, Section 2 deals with segmentation techniques used; Section 3 gives the disparity estimation algorithm used

for analysis, Section 4 with results while analysis and conclusions are given in Section 5.

II. SEGMENTATION ALGORITHMS

For some applications, such as image recognition or stereo vision, whole images cannot be processed, as it not only increases the computational complexity, but it also requires more memory. In literature, number of image segmentation algorithms like K-means [3], mean-shift [4], etc., have been proposed and extensively applied to stereo vision. The algorithm assumes that disparity values vary smoothly in those regions and that depth discontinuities only occur on region boundaries. Purely pixel-based methods are insufficient to express information of the image. The human identifies the objects by analyzing features of the objects such as color, texture and shape. Thus, segmentation-based stereo matching algorithm should be used. Segment-based methods have attracted attention due to their good performance on handling boundaries and texture less regions. They are based on the assumption that the scene structure can be approximated by a set of non overlapping planes in the disparity space and that each plane of target image is coincident with at least one homogeneous color segment in the reference image. Segment based methods perform well in reducing the ambiguity associated with texture less regions and enhancing noise tolerance. The computational complexity is reduced due to much larger segments. Matching becomes much easier even in the presence of noise, intensity variation and slight deviations in segmented area. Noise tolerance is improved by aggregating over pixels with similar colors. One major reason is that small segments may be insufficient for estimating surfaces like slanted planes, while large segments may contain segmentation errors which can affect the accuracy of disparity estimation. Similar colors in image do not always mean similar disparity [1-3]. Mostly, all segment-based stereo matching algorithms employ mean-shift segmentation technique. Taking into consideration all the above points, Particle Swarm Optimization based segmentation algorithms are compared with existing techniques like K-means and Mean Shift Segmentation. Then, disparity estimation was carried out on these segments to do the subjective and objective analysis.

A. Mean Shift Algorithm

Mean shift is a nonparametric iterative algorithm or a nonparametric density gradient estimation using a generalized kernel approach. Mean shift is one of the most powerful clustering techniques [4].

Given n data points $x_i, i = 1, \dots, n$ on a d-dimensional space R^d , the multivariate kernel density estimate obtained with kernel $k(x)$ and window radius h is

$$f(x) = \frac{1}{nh^d} \sum_{i=1}^n k\left(\frac{x-x_i}{h}\right) \quad (1)$$

For radially symmetric kernels, it suffices to define the profile of the kernel $k(x)$ satisfying

$$k(x) = c_{k,d} k(\|x\|) \quad (2)$$

Where $c_{k,d}$ is a normalization constant which assures $k(x)$ integrates to 1. The modes of the density function are located at the zeros of the gradient function $\nabla f(x) = 0$.

$$\nabla f(x) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (x_i - x) g\left(\left\|\frac{x-x_i}{h}\right\|\right) \quad (3)$$

Mean Shift is

$$m(x) = \frac{\sum_{i=1}^n g(x-x_i/h)x_i}{\sum_{i=1}^n g(x-x_i/h)} - x \quad (4)$$

B. K-means Algorithm

K-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem [5]. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. When no point is pending, the first step is completed and an early grouping is done. At this point, we need to re-calculate k new centroids as bary centers of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new centroid. A loop has been generated. As a result of this loop, we may notice that the k centroids change their location step by step until no more changes are done. In other words, centroids do not move any more. Finally, this algorithm aims at minimizing an *objective function*, in this case a squared error function. The objective function

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (5)$$

where $\|x_i^{(j)} - c_j\|$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j , is an indicator of the distance of the n data points from their respective cluster centers. The algorithm is composed of the following steps:

1. Place K points into the space represented by the objects that are being clustered. These points represent initial group centroids.
2. Assign each object to the group that has the closest centroid.
3. When all objects have been assigned, recalculate the positions of the K centroids.
4. Repeat Steps 2 and 3 until the centroids no longer move. This produces a separation of the objects into groups from which the metric to be minimized can be calculated.

C. Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a population based stochastic optimization technique developed by Dr. Eberhart and Dr. Kennedy, inspired by social behavior of bird flocking or fish schooling. PSO is initialized with a group of random particles (solutions) and then searches for optima by updating generations. In each iteration, each particle is updated by following two "best" values. The first one is the best solution (fitness) it has achieved so far. (The fitness value is also stored.) This value is called X_{best} . Another "best" value that is tracked by the particle swarm optimizer is the best value, obtained so far by any particle in the population. This best value is a global best and called g_{best} . After finding the two best values, the particle updates its velocity and positions with following equation (6) and (7).

$$V_R = W * V_R + rand_1 * (r_1 * (X_{best} - X_R)) + rand_2 * (r_2 * (g_{auxR} * g_{best} - X_R)) \quad (6)$$

$$X_R = X_R + V_R \quad (7)$$

where,

W is initial weight,

$r_1 = r_2 = 0.8$,

X_R is randomly generated using maximum and minimum gray level values of image and is updated for N particles g_{aux} is initialized to 1 and

$$g_{best} = \frac{\sum probR[1:X_R(j,1)] * \sum_1^{X_R(j,1)} x_{probR[1:X_R(j,1)]}}{\sum probR[1:X_R(j,1)] - (\sum_1^{Lmax}(probR(1:Lmax))^2)} \quad (8)$$

whose threshold value is kept equal to -10000

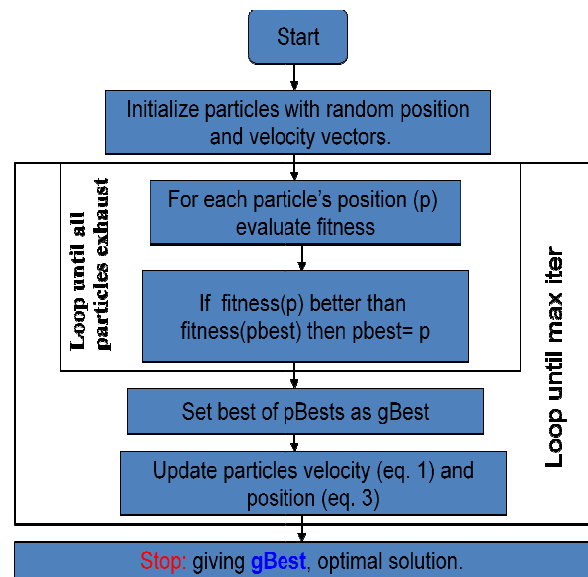


Figure 1. Flowchart of Particle Swarm Optimization

III. DISPARITY ESTIMATION

Disparity refers to the difference in image location of an object seen by the left and right eyes, resulting from the eyes' horizontal separation (parallax). The brain uses disparity to extract depth information from the two-dimensional retinal images in stereopsis. In computer vision,

disparity refers to the difference in coordinates of similar features within two stereo images as given by Bait et.al [8].

In the proposed work, disparity estimation is done in two steps: Disparity Computation and Disparity Optimization. Disparity is computed by finding the cost of matching point $I_l(x,y)$ in the left image to point $I_r(x,y,d)$ in the right image using Sum of Squared Differences (SSD), where matching cost is equal to square of difference of intensity values of pixels at disparity d and can be given as follows.

$$SSD(x, y, d) = \sum_{x,y \in W} (I_l(x, y) - I_r(x, y, d))^2 \quad (9)$$

Disparity is computed as shown in Fig. 2. As a result, we get three sets of disparity cost. Optimization of these 3 sets is done by using winner-take-all method. This method inspects the cost associated with each disparity set via window centered on pixel. Disparity with smallest aggregated cost is selected and given as estimated disparity map.

Disparity estimation was done for four sets of image pair as follows

1. Left and right original images
2. Left and right segmented images using mean shift algorithm
3. Left and right segmented images using K-means
4. Left and right segmented images using PSO

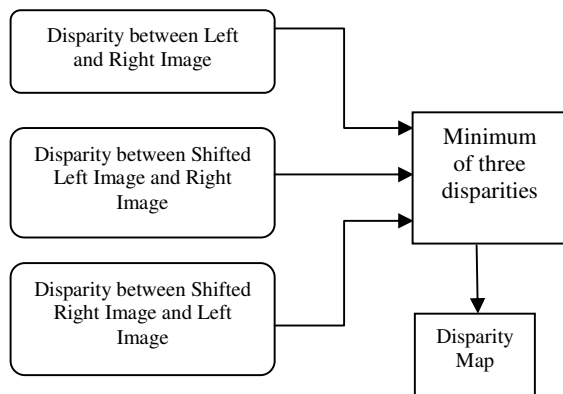


Figure 2. Process of Disparity Estimation

IV. RESULTS

Fig. 3 gives the flow chart for the proposed work for the reconstruction of 3D images. The segmentation stage is skipped when disparity is estimated for the original stereo image pair. It shows that disparity is estimated on the segmented image pair and reconstructed 3D images are also from segmented image pairs. These reconstructed 3D images and disparity obtained are used for both subjective and objective analysis. Fig. 4 illustrates the output of various stages of algorithm for image ‘aloe’, which is one of the image pairs from Middlebury dataset [10].

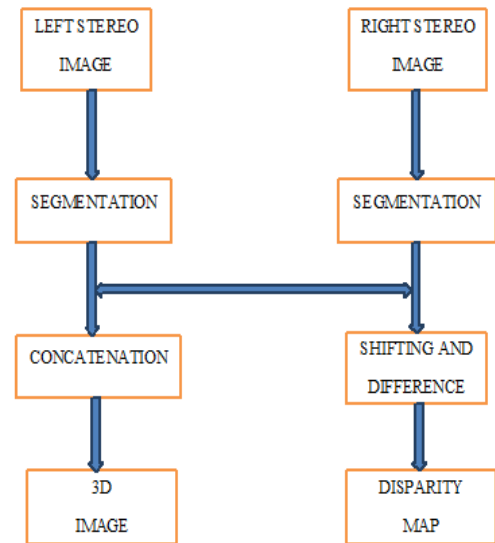


Figure 3. Proposed work flow



Figure. 4(a)



Figure. 4(b)

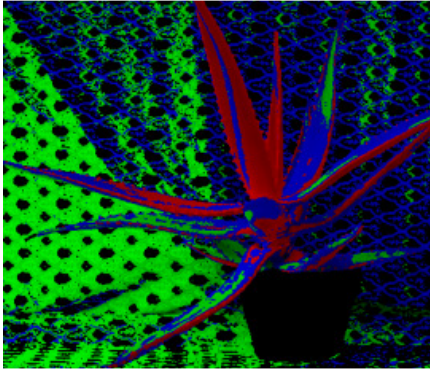


Figure. 4(c)

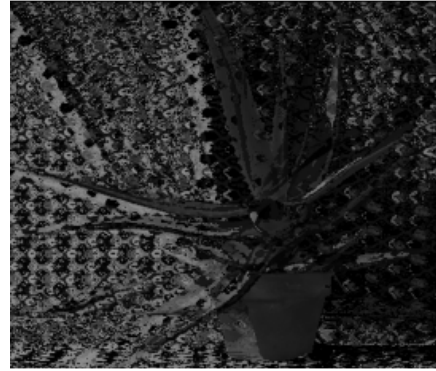


Figure. 4(g)

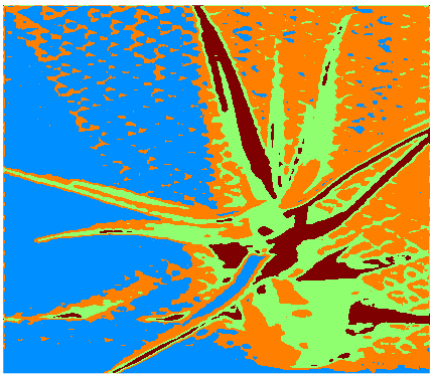


Figure. 4(d)

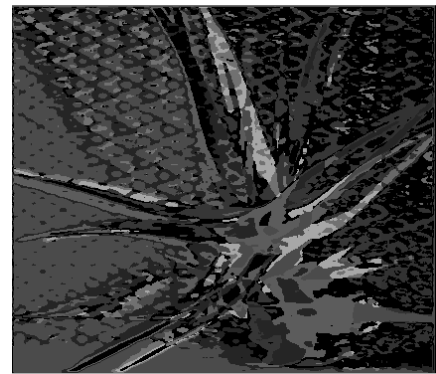


Figure. 4(h)

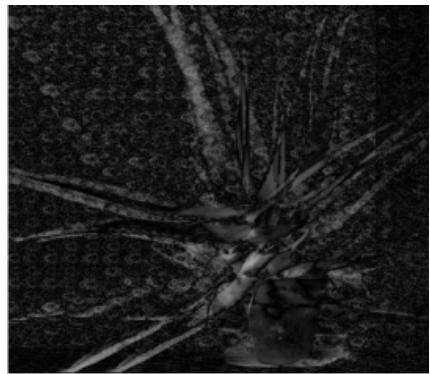


Figure. 4(e)

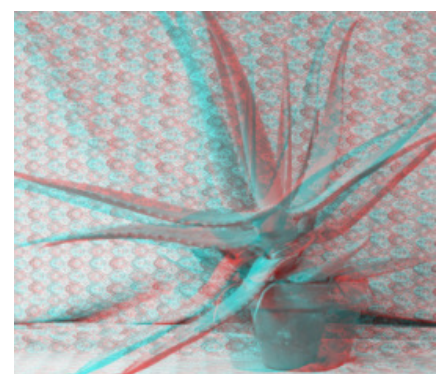


Figure. 4(i)

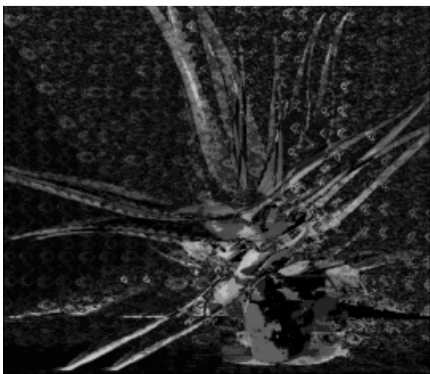


Figure. 4(f)

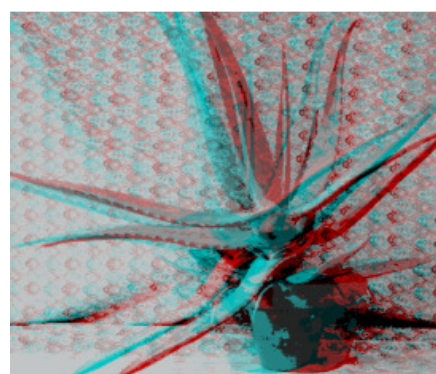


Figure. 4(j)

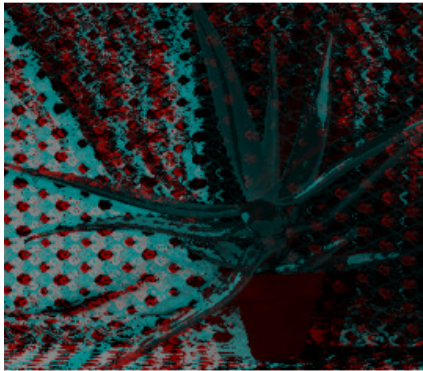


Figure. 4(k)

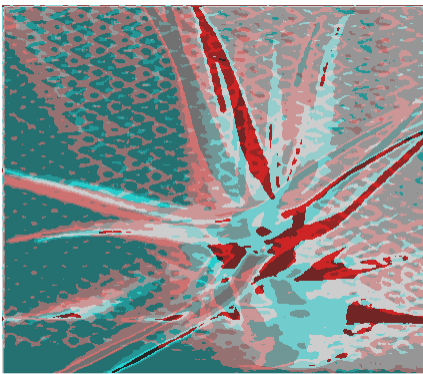


Figure. 4(l)

Figure 4. Implementation Results on 'Aloe' image

Fig. 4(a) represents the original left view from a stereo image pair. Fig. 4(b) is the segmented left image using PSO, Fig. 4(c) is the segmented left image using K-means, and Fig. 4(d) is the segmented left image using Mean Shift algorithm. Similar segmented images were obtained for the right view of stereo pair. Original stereo pair and segmented images pair were given as input for disparity estimation one-by one.

Fig. 4(e) is disparity obtained using original stereo pair while Fig. 4(f), Fig. 4(g) and Fig. 4(h) are disparities obtained from segmented images, using PSO, K-means Algorithm and Mean Shift Algorithm respectively.

Fig. 4(i) is the reconstructed 3D image obtained from original image pair. Fig. 4(j), Fig. 4(k) and Fig. 4(l) are the reconstructed 3D images obtained from PSO algorithm, K-means algorithm and Mean Shift algorithm, respectively. These images were tested on 100 subjects for subjective analysis.

V. CONCLUSION AND FUTURE WORK

All the algorithms were implemented on Middlebury database and were compared for the performance parameters like PSNR, compression ratio and number of depth levels extracted. Performance parameters for 12 such images are given here for reference.

Fig. 5 gives comparison for PSNR of disparity estimation using various segmentation techniques where

vertical axis represents the PSNR in db with respect to disparity estimation of original image. It was observed that PSNR for all the three segmentation techniques is almost same.

Fig. 6 shows the plot of percentage compression on its vertical axis for different images. The compression ratio for PSO was calculated on an average of 50 % and for K-means was calculated on an average of 57% for the Middlebury database [10]. Even though the compression ratio of reconstructed 3D image based on PSO segmentation technique is less as compared to K-means, it can be seen that the subjective quality of PSO based 3D reconstructed images, (Fig. 4(j), Fig. 4(k) and Fig. 4(l)), gives much better vision than K-means and mean shift algorithms, which was found to 85% good and comparable to the 3D reconstruction using original images.

Fig. 7 shows the comparison of number of estimated depth levels for different algorithms. Number of Depth levels estimated using PSO segmentation is almost same to the number of depth levels estimated from original image. This may be one of the reasons why the subjective analysis for 3D image reconstructed using PSO segmentation provides better results compared to other two segmentation techniques.

PSO algorithm retains the original information of the image even after segmentation. PSO based segmented images provide better disparity estimation, with good number of estimated depth levels. The subjective quality of 3D images obtained using PSO technique is also better. In future, it can be one of the valuable segmentation techniques for color stereo vision. This paper also laid a good foundation for the study of improvement in the compression ratio using techniques like Functional Swarm Optimization (FSO) and Darwinian PSO (DPSO). The proposed algorithms not only approximate the boundaries of interest but may reduce the Computational complexity.

REFERENCES

- [1] Z. Wang and Z. Zheng, "Region Based stereo Matching Algorithm Using Cooperative Optimization", Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, 23-28 June 2008, Page(s): 1 – 8
- [2] V. Borisagar and M. Zaveri, "A Novel Segment-based Stereo Matching Algorithm for Disparity Map Generation", 2011 International Conference on Computer and Software Modelling IPCSIT vol.14 (2011) © (2011) IACSIT Press, Singapore.
- [3] A. Klaus, M. Sormann, and Karner, "K-Segment-Based Stereo Matching Using Belief Propagation" Pattern Recognition, 2006. ICPR 2006. 18th International Conference, Page(s):15-18
- [4] D. Comaniciu and P. Meer, "Mean shift: A Robust Approach Toward Feature Space Analysis", Pattern Analysis and Machine Intelligence, IEEE Transactions 2002, Page(s):603-619
- [5] K. Javed, "The Behaviour of k-Means: An Empirical Study", Electrical Engineering 2008. ICEE 2008, Second International Conference on 25-26 March, Page(s):1-6
- [6] P. Ghamisi, M. Couceiro, J. Benediktsson, and M. Ferreira, "An Efficient Method for segmentation of Images based on Fractional

- Calculus and Natural Selection Expert Systems with Applications”, An International Journal, Volume 39, Issue 16, November 2012.
- [7] F. Bergh and A. Engelbrecht, “A study of Particle Swarm Optimization Particle Trajectories”, Information Sciences: An International Journal, Volume 176, Issue 8, April 2006.
- [8] T. Bait, D. Boudaoud, B. Matsuzewski, and L. Shark, “An Efficient Feature Based Matching Algorithm for Stereo Images”, Geometric Modelling and Imaging, 2006, Page(s):195-202
- [9] J. Joglekar and S. Gedam, “Area Based Stereo Image Matching Technique Using Hausdorff Distance And Texture Analysis”, International Journal on Emerging Technology and Advanced Engineering, ISSN 2250-2459, Volume 2, Issue 1, January 2012.
- [10] www.vision.middlebury.edu/stereo/dataset, date: 23rd Feb 2013

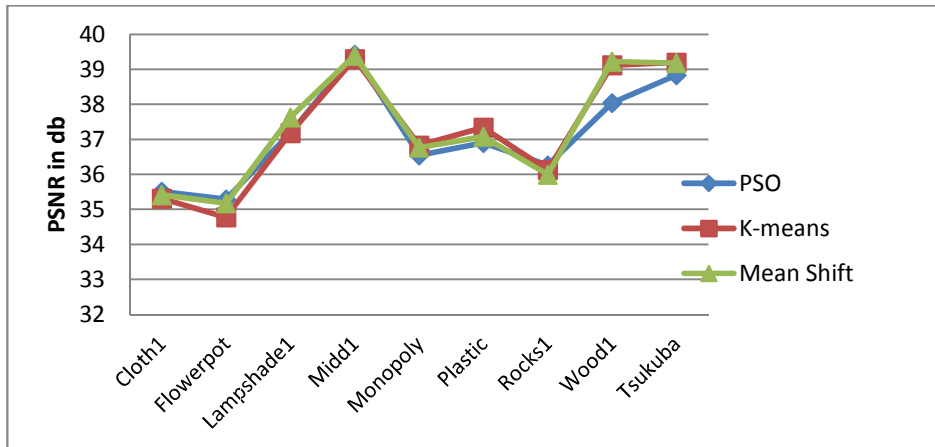


Figure 5. Plot of PSNR of three segmentation techniques

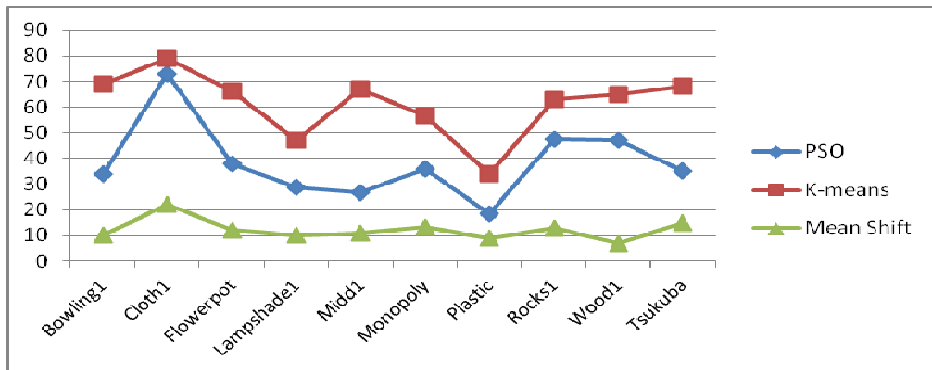


Figure 6. Plot of % compression Ratio of three segmentation techniques

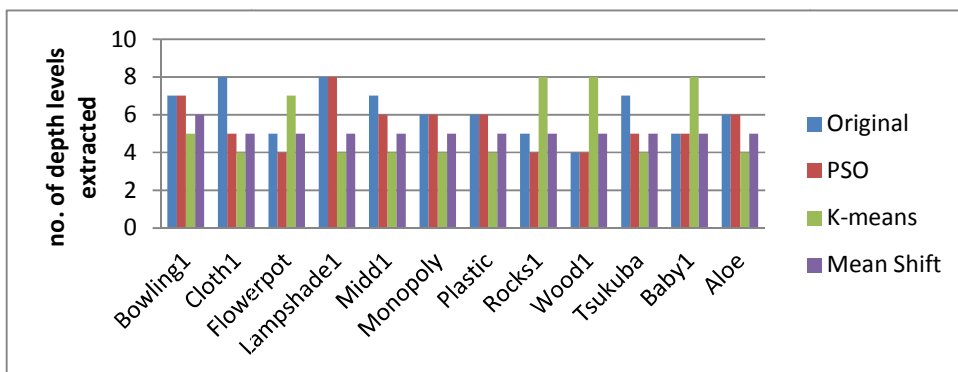


Figure 7. Plot of estimated depth levels of three segmentation techniques